

Article

Bayesian Neural Networks for Thermal Resilience Optimization Under Future Climate Scenarios: A Case Study of Affordable Housing in Tropical Regions

Ibrahim Elwy ^{1,*} , Yasser Ibrahim ² , Fatima Zahrau Muhammed ¹ , Xiong Zhilun ¹ and Aya Hagishima ^{1,3} 

¹ Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Kasuga City 816-8580, Japan; fatyazhra93@kyudai.jp (F.Z.M.); xiong.zhilun.374@s.kyushu-u.ac.jp (X.Z.); ayahagishima@kyudai.jp (A.H.)

² Architectural Engineering Department, Military Technical College, Cairo 11766, Egypt; yasser_ibrahim@mtc.edu.eg

³ Faculty of Engineering Sciences, Kyushu University, Fukuoka 819-0395, Japan

* Correspondence: ibrahim.elwy@kyudai.jp

Abstract

Global warming and increasing heat events necessitate long-term assessments of passive design strategies to ensure thermal resilience under future climatic conditions. Although machine-learning-based Surrogate Models (SMs) offer timely approximation of building performance compared to conventional simulation-based approaches, the lack of uncertainty quantification raises concerns about the reliability of their design optimization outcomes. This study aims to develop a robust surrogate-assisted optimization framework, based on a probabilistic Bayesian Neural Network (BNN) model and supported by an uncertainty-aware objective function. The framework is applied to an affordable housing case study in Surakarta, Indonesia, evaluating its generalizability under current and future climatic scenarios for 2050, 2070, and 2090. Thermal resilience is assessed through overheating hours exceeding acceptability limits in Southeast Asian context, using a parametric workflow implemented in Ladybug-tools and Grasshopper 3D. Compared to simulated test data, the BNN model demonstrates reliable predictive accuracy and probabilistic inference ($R^2 = 0.99$, MAE = 0.52%, CRPS = 0.38%). Furthermore, validation against re-evaluated optimal solutions shows low error ranges (RMSE = 0.43%, MAE = 0.33%), outperforming the deterministic SM optimization approach—using Artificial Neural Networks—by a factor of five. Overall, the uncertainty-aware framework provides a feasible, overconfidence-resistant, and reliable surrogate-assisted optimization method, identifying optimal solutions closely matching those from simulation-based optimization while reducing computational time by 96%.

Keywords: Bayesian Neural Networks; Probabilistic Surrogate Models; thermal resilience; affordable housing; uncertainty-aware optimization; climate change



Received: 8 December 2025

Revised: 4 January 2026

Accepted: 7 January 2026

Published: 13 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Global urbanization is projected to reach 70% by 2050 [1], with nearly 90% of this growth occurring in developing countries [2]. Tropical regions, which accommodate 43% of the world's population [3], are experiencing rapid and often unplanned urban expansions, exacerbating social and spatial inequalities that increase the vulnerability of low-income communities to climate change impacts [4]. This vulnerability is predominantly driven by rising housing and operational costs—particularly energy expenses [5]—together with the

poor thermal performance of affordable dwellings constructed using lightweight materials, minimal insulation, and limited provision for natural ventilation or active cooling [6–9]. These characteristics contribute to severe indoor overheating under global warming and intensifying extreme events [10–13], underscoring the urgent need for climate-responsive passive design strategies as long-term adaptation measures.

The short-term effectiveness of passive design strategies for affordable housing in tropical regions has been extensively documented [6,14–17], particularly the roles of natural ventilation [18], shading [19], and insulation [20] in mitigating overheating risk and reducing energy demand. However, long-term assessments of building performance (BP) under projected—rather than present—climatic conditions remain comparatively limited. Existing studies addressing future climate impacts (e.g., Ref. [21–24]) often evaluate a narrow set of climate scenarios or investigate a restricted subset of passive design parameters. Furthermore, while recent analyses have demonstrated that the effectiveness of specific passive measures may degrade under future climate projections [23–26], systematic investigations that quantify their relative and evolving influence on thermal resilience across multiple climate change scenarios are still scarce. This highlights the need for unified, large-scale evaluation frameworks capable of supporting reliable design decisions across multiple future climate scenarios, at a scale that is computationally challenging using physics-based simulations alone.

To this end, machine-learning-based Surrogate Models (SMs) have gained increasing attention as effective and computationally efficient BP approximation methods [27,28]. SMs emulate the behavior of complex physics-based simulations in real-time, facilitating design space exploration, optimization, and sensitivity and uncertainty analysis [29]. Within the field of Building Performance Optimization (BPO), however, the majority of surrogate-assisted frameworks remain fundamentally deterministic, producing single-point predictions without considering surrogate uncertainty [28]. This limitation becomes particularly critical when optimization algorithms operate in sparsely sampled or under-represented regions of the design space, where surrogate predictions can be unreliable yet overconfident, misleading the search process toward false optima [28]. This risk is further amplified in generalized SMs trained across multiple climates or building contexts, where prediction conditions may substantially deviate from those represented in the original training data.

Despite these recognized limitations, existing surrogate-assisted optimization frameworks largely lack mechanisms to quantify and incorporate surrogate uncertainty into design decision-making. In practice, optimization algorithms implicitly assume uniform confidence across the design space, treating surrogate predictions as equally reliable regardless of training data density, interpolation distance, or contextual variability. As a result, current frameworks may yield design solutions that appear optimal in a deterministic sense but cannot be considered robust or reliable when SM predictive uncertainty is non-negligible.

Table 1 summarizes recent studies that incorporated SMs for BP prediction and optimization, with emphasis on surrogate uncertainty treatment. Advances in probabilistic SMs based on Gaussian Process (GP) and Bayesian Deep Learning (BDL) have emerged as promising tools for annual energy prediction [30], hourly load forecasting [31], and uncertainty-guided infill (adaptive) sampling for efficient optimization [32]. Nevertheless, most probabilistic approaches in building performance research either analyze uncertainty solely for predictive interpretation (Category B) or exploit predictive variance as a heuristic for sampling efficiency (Category C). In such cases, surrogate uncertainty is not explicitly integrated into the optimization objective or decision criteria, particularly when assessing thermal resilience across multiple climate scenarios. Even when Ensemble Learning

Methods (ELMs) are employed to enhance optimization reliability (e.g., [33]), the resulting improvements primarily stem from approximating variability among base models rather than formally quantifying surrogate uncertainty. Consequently, the optimization outcomes of these methods remain effectively deterministic and still exhibit overconfidence in sparsely sampled or noisy regions of the design space.

Table 1. Surrogate models and uncertainty treatment in BP prediction and optimization.

Category *	[Refs.]	SM Type	Application Context	Uncertainty Treatment	Use in Design Optimization/Exploration	Relevance to This Work
A	[33,34]	ELM	Single- or multi-objective optimization, minimizing thermal comfort or carbon emissions	None	Ensemble methods to improve BP prediction accuracy	Deterministic SM-assisted BPO with no surrogate uncertainty quantification
	[35]	ELM	Optimizing annual energy consumption	Ensemble disagreement (committee variance)	Adaptive (infill) sampling	Deterministic SM-based BPO with uncertainty-guided sampling
B	[36]	BDL	Energy flexibility assessment	Epistemic + aleatoric uncertainty	BP prediction/forecasting only (no optimization)	- Probabilistic SMs in building performance research - Uncertainty is analyzed/visualized, but not used for design optimization
	[30]	BDL, GP	Annual load prediction	Epistemic uncertainty		
	[31]	GP	Time-series load forecasting	Explicit GP predictive distribution		
	[37]	GP	Calibration of energy models	Explicit Bayesian posterior		
C	[32,38–40]	GP	Single- or multi-objective optimization, minimizing energy consumption, carbon emissions, and cost	Implicit GP predictive variance (epistemic uncertainty)	Efficient Global Optimization (EGO)/Evolutionary search for infill sampling	- SM-assisted BPO - Uncertainty guided efficient exploration by maximizing Expected Improvement (EI) - Deterministic decision making, no uncertainty-aware optimization
[This study]		BDL	Multi-scenario thermal resilience optimization under climate change	Explicit aleatoric + epistemic uncertainty (posterior sampling)	- A novel, resilience-oriented BPO framework using a generalized BNN - Enabling surrogate uncertainty-informed building design decision - Uncertainty-aware objective using risk metrics with epistemic and coverage penalties	

* Category (A): Classical deterministic SM-assisted BPO; (B): Probabilistic SMs for other applications (no optimization); and (C): Probabilistic SMs for optimization.

To address this gap, this study is the first to formulate surrogate-assisted thermal resilience optimization as a probabilistic decision-making framework, in which surrogate uncertainty is explicitly incorporated into the objective function. A generalized Bayesian Neural Network (BNN) surrogate model is developed to predict thermal discomfort while quantifying surrogate uncertainty through predictive distributions. Apart from physical building model uncertainty, the framework distinguishes between aleatoric uncertainty, representing inherent variability and noise in the simulated data, and epistemic uncertainty, reflecting the BNN's lack of knowledge due to uneven or limited sampling of the design space.

The proposed framework is demonstrated using a representative affordable housing case study in Surakarta, Indonesia, evaluated under current and future climatic scenarios for 2050, 2070, and 2090. Thermal discomfort is adopted as a resilience-oriented performance metric, enabling the investigation of the effectiveness of passive design strategies under

progressive climate change. By integrating predictive distributions, epistemic uncertainty penalties, and coverage-aware risk metrics into the optimization objective, the framework enables robust design decision-making while maintaining computational efficiency.

In summary, this study aims to:

- Develop and validate a probabilistic generalized SM capable of simultaneously predicting thermal discomfort and quantifying surrogate uncertainty;
- Systematically investigate the changing influence of passive design strategies on thermal resilience across future climate change scenarios;
- Demonstrate robust surrogate-assisted BPO, enabling reliable design decisions informed by surrogate predictive uncertainty.

2. Methodology

2.1. Case Study, Energy Modeling and Climatic Scenarios

Based on extensive field measurements and observations of the indoor environmental condition and occupant behavior in three slum-upgrading projects in Surakarta, Indonesia [41], a northwest-oriented two-story dwelling (T1) was selected as the case study. The dwelling (Figure 1) is constructed using the dry-joint precast concrete system “Rumah Unggul Sistem Panel Instan” (RUSPIN), which represents a widely adopted affordable housing typology in Indonesia, particularly in post-disaster reconstruction due to its low-cost, rapid assembly, and seismic resistance [42,43]. The house comprises a living space and a bathroom on the first floor and a bedroom on the second floor with an adjacent terrace, often used as a second bedroom.

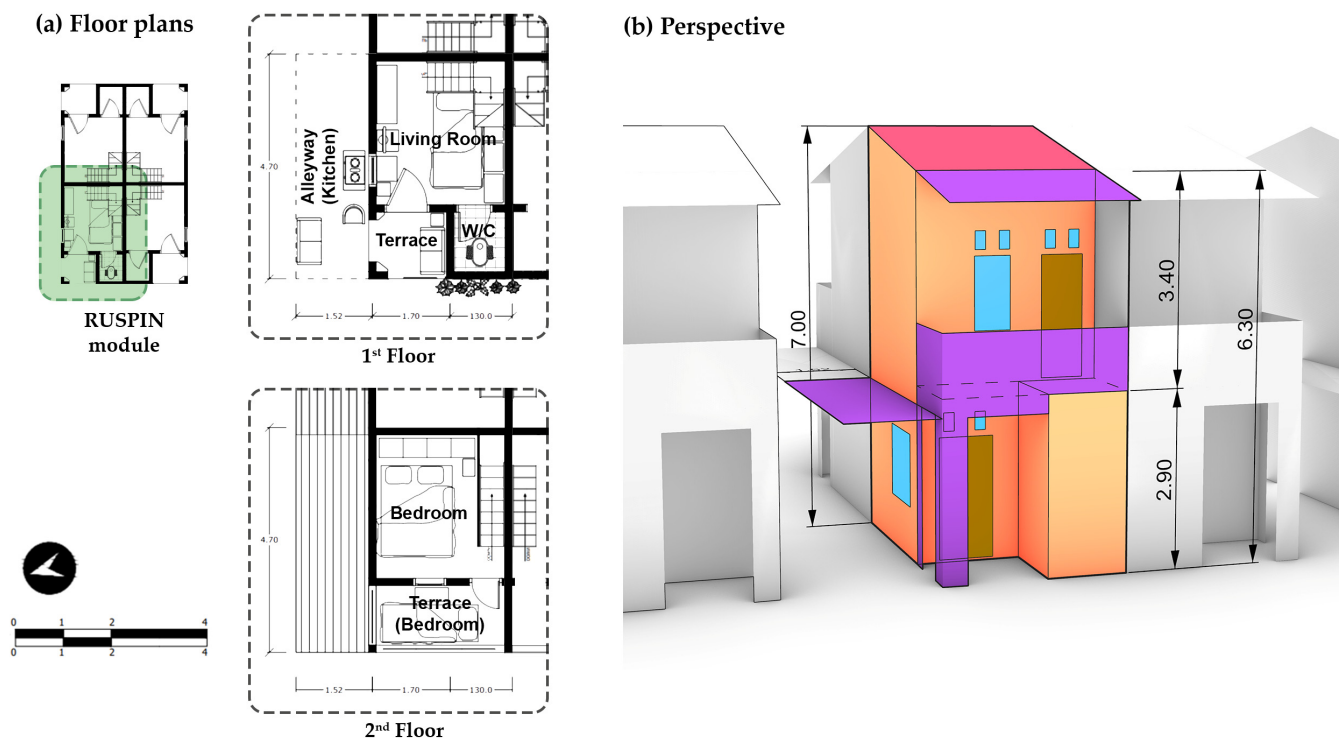


Figure 1. Floor plans (a) and Honeybee model (b) for the basic configuration of the target dwelling T1, adopted from [18].

A detailed building energy model was developed using a parametric simulation workflow implemented through the Ladybug and Honeybee plugins of Grasshopper 3D [44]. The EnergyPlus model strictly reflects the as-built geometry, spatial configuration, construction materials, and documented occupant behavior of the dwelling. It was modeled as three

separate thermal zones—living room, bathroom, and bedroom, whereas the surrounding buildings were modeled as opaque shading objects. Envelope material properties are summarized in Table 2. Internal partition walls shared with the adjacent dwelling were modeled as adiabatic surfaces, while the ground was represented as deep moist soil.

Table 2. Physical and thermal properties of building envelope materials used in simulation.

Building Component	Material	Thickness [m]	Density [kg/m ³]	Thermal Conductivity [W/m·K]	Specific Heat [J/kg·K]
Walls	Autoclaved Aerated concrete blocks	0.15	1400	0.6	840
Roof	Galvanized steel plate	0.0015	7680	44.98	418
Windows	Clear single glass	0.006	2500	0.9	750
Entrance door	Wood	0.001	608	0.08	1200
Floor	Ceramic porcelain	0.01	2300	1.3	840
	Concrete	0.10	2322	2.31	832

Airflow interaction between floors was represented using the default configuration of Honeybee Airflow Network (AFN) component, with the staircase opening modeled as an air boundary to account for the wind- and buoyancy-driven air exchange. In addition, door ventilation settings were customized using a Python script, which is combined with an energy management system (EMS) for aperture operation schedules, ensuring consistency with monitored occupancy patterns and ventilation behavior. To ensure reproducibility, details regarding customized scripts for door operability, EMS, and operation schedules are provided in Supplementary Material S1.

Ventilation-related parameters and material properties were calibrated based on the available field measurements and in situ observations reported in [41]. Following calibration, the simulation model was verified against measured indoor air temperatures over a 20-day period (7–26 October 2023). Since the bedroom zone exhibited greater overheating vulnerability from direct exposure to untreated metal roofing, it was selected for examining the performance impact of passive design interventions. Verification results showed a consistent performance between simulated and measured temperatures in the bedroom zone, with a Root Mean Squared Error (RMSE) of 1.0 °C, and a Mean Absolute Error (MAE) of 0.8 °C. In addition, normalized Mean Bias Error (nMBE) and Coefficient of Variation in the RMSE (CVRMSE) were found to be +0.14 and 3.15, respectively, satisfying frequently adopted calibration criteria of hourly data, as reviewed by Coakley et al. [45].

Although the monitoring period was limited in duration, it corresponds to typical climatic conditions in Surakarta, which exhibits relatively small seasonal variability in ambient air temperature (Figure 2). The purpose of the calibrated energy model is not to predict absolute long-term indoor temperature trajectories, but to provide an empirically informed baseline for comparing thermal comfort performance of alternative passive design strategies. Accordingly, the calibrated parameters—of envelope material properties and ventilation-related behavior—were considered representative for establishing the base case model used in subsequent parametric simulations. Long-term material aging, climatic variability, or stochastic occupant behavior were therefore not considered by the calibration process. Addressing these additional sources of uncertainty is beyond the scope of the present study, which focuses on the robustness of surrogate model predictions in design optimization applications under climate variability.

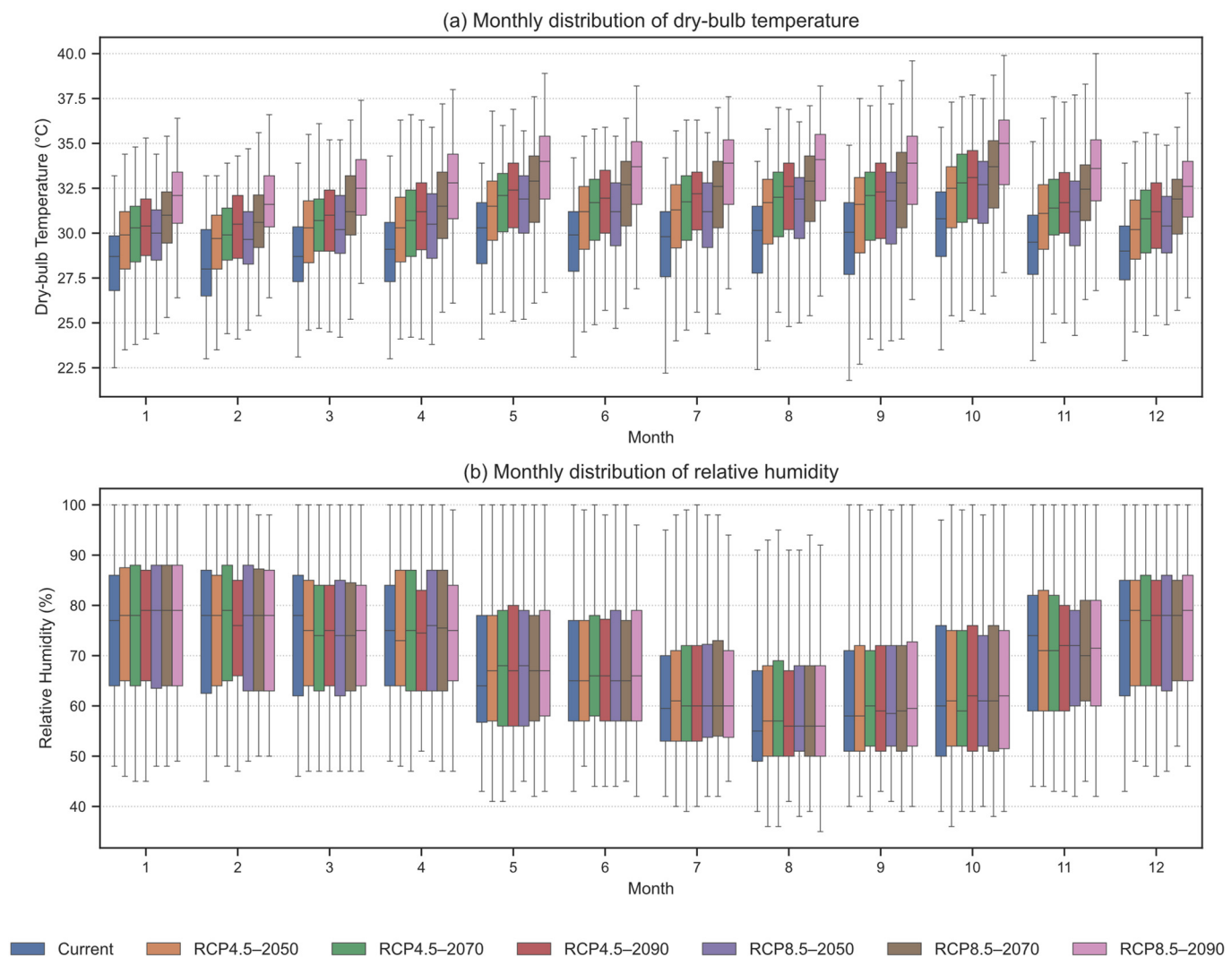


Figure 2. Monthly distribution of dry-bulb temperature (a) and relative humidity (b) of the seven climate scenarios.

The energy simulation boundary conditions were generated by Meteonorm v8.2 [46] as Typical Meteorological Year (TMY) datasets of Surakarta, representing the current (contemporary) climate and six future weather files for the years 2050, 2070, and 2090, each under Representative Concentration Pathways (RCPs) of 4.5 and 8.5—indicating medium and high emission scenarios, respectively. The dry-bulb temperature distributions for the seven meteorological datasets (Figure 2a) show a pronounced warming trend across all months. Moreover, the corresponding relative humidity distributions are displayed in Figure 2b, revealing seasonal variation while exhibiting minor shifts among temporal datasets. Further details on the generation of future climate scenarios, including the underlying climate models ensemble, scenario assumptions, and the final generation of hourly TMY files, are provided in Supplementary Material S2.

2.2. Database Generation and Preprocessing

2.2.1. Passive Design Parameters

We adopted nine design parameters (p2–p10) for database generation of the bedroom model, representing envelope- and geometry-related passive cooling strategies that are widely investigated in affordable housing in tropical regions. Room orientation and climatic scenario (weather file) represented two additional parameters (p1 and p11), to account for contextual variability. The name and acronym of each parameter, as well as their upper and

lower bounds, are shown in Table 3. The lower bounds of zero in p4, p5, and p6 indicate the absence of the corresponding design strategy. Moreover, maximum values of WWRs were constrained by the available façade area, considering the prefabricated structural system and the operable terrace door in the front façade, as illustrated in Figure 3.

Table 3. Passive design parameters used for database generation.

Parameter	Name	Units	Bounds
p1	Orientation (Or)	° (degrees)	0–360
p2	WWR for the front window (WWR-Fr)	%	15–55
p3	WWR for the side window (WWR-Si)	%	20–75
p4	Overhang shading projection (Shd_Proj)	m	0.00–1.50
p5	XPS insulation thickness for roof (R_XPS_t)	m	0.00–0.15
p6	XPS insulation thickness for walls (W_XPS_t)	m	0.00–0.15
p7	Solar absorptance for roof (R_SA)	–	0.10–0.90
p8	Solar absorptance for wall (W_SA)	–	0.10–0.90
p9	Windows thermal transmittance (Win_U)	kWh/m ² K	0.11–1.98
p10	Windows solar heat gain coefficient (SHGC)	–	0.11–0.77
p11	Weather file index	–	0–6

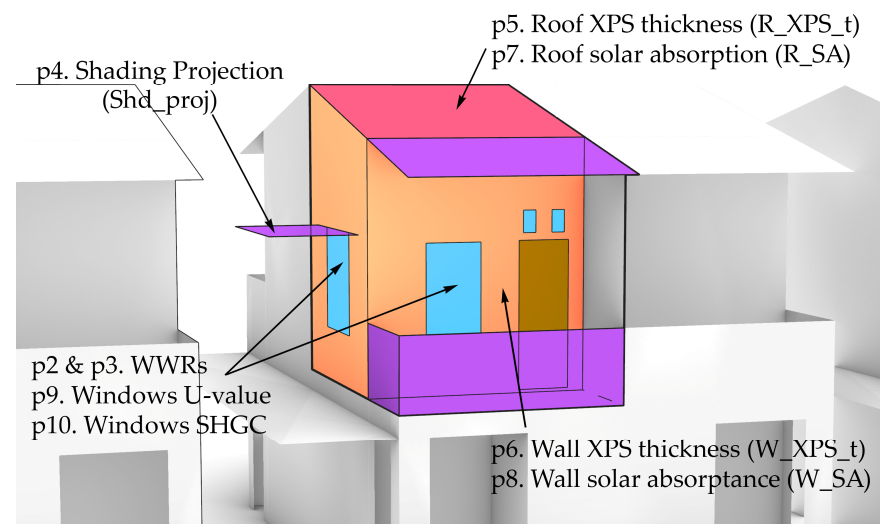


Figure 3. Passive design parameters (p2–p10) used for the bedroom zone.

Operational strategies related to occupant behavior were not treated as design parameters in this study. Instead, they were prescribed based on field observations of the target dwelling and held constant across all design cases and climate scenarios. As detailed in Supplementary Material S1, entrance door operation followed fixed schedules reflecting observed practices that are unrelated to thermal discomfort, while window openings were assumed to operate at their maximum feasible availability throughout the year. This assumption is consistent with typical behavior in naturally ventilated low-income housing in similar climates, where occupants extensively utilize available openings for ventilative cooling under current conditions and where projections of future behavioral change remain highly uncertain. By fixing operational behavior, the present study isolates the performance and resilience of passive design strategies under progressive climate change, acknowledging that scenario-dependent operational behavior represents an additional source of uncertainty beyond the scope of this work. Occupant adaptation to outdoor climate is therefore represented at the population level through the adopted adaptive comfort formulation with wider acceptability limits, as discussed in Section 2.2.2.

2.2.2. Indoor Thermal Comfort Metric

Previous investigations in comparable climatic contexts have reported higher thermal adaptive capacity than that prescribed by the ASHRAE standard 55 Adaptive Comfort Model (ACM) [47,48]. Accordingly, we adopted an ACM formulation specifically tailored for free running and mixed-mode buildings in Asian and tropical climates, using the comfort Equation in (1) derived by [49,50] as a regional adjustment to the original framework proposed by de Dear et al. [51].

$$T_{o,comf} = 0.31 T_{po} + 19.1 \quad (1)$$

where $T_{o,comf}$ is the comfortable indoor operative temperature, and T_{po} is the prevailing mean outdoor temperature. Within the ACM framework, acceptability limits of ± 2.5 °C and ± 3.5 °C around $T_{o,comf}$ correspond to 90% and 80% occupant acceptability limits, respectively. The wider 80% acceptability band reflects a greater degree of thermal tolerance, which has been shown to better represent adaptive behavior in naturally ventilated and mixed-mode buildings across various contexts [52,53].

In this study, we assessed thermal comfort using the 80% acceptability limits, where thermal discomfort was quantified as the percentage of annual hours exceeding these limits. Under extreme future scenarios, such as RCP8.5–2090, the maximum outdoor dry-bulb temperature exceeds 38 °C (Figure 2a), substantially diminishing the effectiveness of natural ventilation and challenging occupants' adaptive capacity. This highlights the need for robust passive cooling strategies to ensure thermal resilience. Since overheating was most pronounced in the bedroom zone [54], we adopted the annual percentage of hours hotter than the acceptability limit (BR_pct_hot) as the thermal comfort metric.

Notably, the absolute magnitude of pct_hot is inherently dependent on the selected comfort model and the width of the acceptability band. Alternative ACM formulations or narrower comfort ranges would lead to completely different discomfort levels. However, the relative comparison of passive design alternatives and the observed trends progressive climate change should remain consistent when applying the same ACM and acceptability criterion for all design cases.

2.2.3. Sampling Strategy and Data Preprocessing

A previous study evaluated the generalizability of SMs across multiple climatic conditions by systematically comparing the settings of ELMs and their sampling strategies [34]. The results indicated that increasing the sample size beyond 500 cases per climatic condition significantly raised computational cost but provided only marginal improvements in SM accuracy. Therefore, 500 Sobol' sequence (SS) samples were generated for each of the seven climatic scenarios, resulting in a total of 3500 samples in the present study.

It should be noted that the SM is intended to emulate the response of the calibrated physics-based simulation model for large-scale parametric applications with multiple design variables—such as sensitivity analysis and optimization—rather than to predict measured indoor conditions directly. Therefore, the generated database of passive design parameters and simulation output was subsequently used for conducting all SM training, testing, and validation, following common practice in surrogate-assisted BPO studies.

Data preprocessing was subsequently applied to transform raw parameters into trainable features through feature adjustment and normalization. The adjustment included encoding the orientation parameter (p1) into sine and cosine features (x1 and x2 in Table 4) to account for its periodic nature. Additionally, several continuous climatic indicators—namely annual dry-bulb temperature (T_{db}), relative humidity (RH), global horizontal radiation (GHI), and wind speed (V)—were initially considered based on previous cross-

climatic surrogate modeling studies [55,56], where their suitability to replace categorical weather data indices (p11) was investigated.

Table 4. SM input and output features.

Feature	Parameter	Name	Units	Bounds
x1	p1	Or_sin	–	0–1
x2	p1	Or_cos	–	0–1
x3–x11		Same as p2–p10 in Table 3, respectively		
x12	p11	CDD	–	1758.82–3142.04
Output (y)		BR_pct_hot	%	0–100

Although these meteorological variables are critical drivers of thermal comfort and natural ventilation, their effects are already fully captured in the underlying EnergyPlus simulations used to generate the training dataset. At the surrogate modeling stage, climatic indicators were therefore introduced solely to distinguish temporal climatic scenarios, rather than to reconstruct meteorological conditions. An exploratory statistical analysis of the seven weather datasets (Table 5) shows that annual distributions of key meteorological variables exhibit nearly invariant distributions across scenarios, with changes in mean (μ), standard deviation (σ), and percentile ranges (p10 and p90) remaining limited and without clear separation between both RCP scenarios.

Table 5. Annual statistics of key meteorological parameters across the seven climate scenarios.

Weather Dataset	T_db				RH				GHI				V				CDD *	HDD *
	μ	σ	p10	p90	μ	σ	p10	p90	μ	σ	p10	p90	μ	σ	p10	p90		
Current	29.18	2.54	25.5	32.3	68.8	14.91	50	89	395.7	292.8	9	822	1.85	1.79	0.2	4.2	1758.82	0
RCP4.5–2050	30.55	2.57	26.8	33.5	69.31	14.64	51	89	405.9	295.4	10	825	1.85	1.79	0.2	4.2	2255.79	0
RCP4.5–2070	30.99	2.58	27.2	34	69.4	14.71	51	90	408.2	296.7	10	831	1.85	1.78	0.2	4.2	2412.39	0
RCP4.5–2090	31.34	2.58	27.6	34.4	69.4	14.58	51	90	409.5	294.9	9	828	1.86	1.79	0.2	4.2	2544.88	0
RCP8.5–2050	30.77	2.58	27.1	33.9	69.48	14.87	51	90	400.8	297.3	9	829	1.85	1.8	0.2	4.2	2339.5	0
RCP8.5–2070	31.84	2.59	28.1	34.9	69.53	14.79	51	90	401.4	295.4	9	829	1.84	1.79	0.2	4.1	2729.93	0
RCP8.5–2090	32.98	2.61	29.26	36.1	69.59	14.79	51	90	402.4	296.9	9	827	1.85	1.8	0.2	4.2	3142.04	0

* CDD and HDD were, respectively, calculated using default base temperatures of 23 °C and 18 °C.

In contrast, the Cooling Degree Days (CDD) indicator provides a compact and physically meaningful feature that directly reflects the long-term climate-change effects relevant to overheating risk. Low varying climate indicators—including Heating Degree Days (HDD)—were consistently negligible for the study location and therefore provided no additional explainability across scenarios. Accordingly, CDD (x12) was retained as the sole climatic feature used to represent temporal climate variability in SM training.

Moreover, all input features were normalized using standard scaler from Scikit-learn Python library [57], to align the distribution of feature magnitudes of mismatching scales with the prior assumptions of model weights, leading to smoother and more stable training. Finally, the full dataset of twelve input features and a single output of bedroom's pct_hot (Table 4) was partitioned into 70% for training, and 30% equally allocated to validation and testing. The partitions were stratified by CDD, to ensure a balanced sample distribution across all climate scenarios.

2.3. Probabilistic Surrogate Model Development and Validation

We developed a Bayesian Neural Network (BNN) model, using TensorFlow Probability [58], to achieve prediction accuracy while quantifying the underlying uncertainty of

model outputs through probabilistic distributions of model parameters. Unlike deterministic Artificial Neural Networks (ANN) that learn fixed weights, the BNN parameters are characterized by random probability distributions that evolve during the training process, allowing the model to express the degree of uncertainty of its predictions.

The network architecture consisted of two variational hidden layers with 128 and 64 neurons, respectively, both employing the Rectified Linear Unit (ReLU) activation. These variational layers (DenseVariational) learned probability distributions—rather than point estimates—over weights and biases. The output layer parametrized a univariate normal likelihood using two neurons, the predictive mean (μ) and standard deviation (σ), while a final DistributionLambda layer defined the probability distribution for the target variable (BR_pct_hot), enabling probabilistic inference rather than deterministic regression. Generally, the Bayesian learning process aims to estimate the posterior distribution $p(\theta|D)$ of model parameters (θ), given training data (D), as expressed in Equation (2):

$$p(\theta|D) \propto p(D|\theta) p(\theta) \quad (2)$$

where $p(\theta)$ represents the prior distribution of BNN weights and biases; $p(D|\theta)$ is the likelihood of observed data given parameters; and $p(\theta|D)$ represents the posterior distribution approximated via variational inference.

We assigned a trainable normal prior for the weights and biases for each DenseVariational layer, and defined an approximate posterior as a mean-field Gaussian distribution with learnable mean and scale parameters. To optimize these distributions, the BNN minimized the negative Evidence Lower Bound (-ELBO), which defined the model's overall loss function $\mathcal{L}(\theta)$, as expressed in Equation (3):

$$\mathcal{L}(\theta) = -\mathbb{E}_{q(\theta)}[\log p(D|\theta)] + \text{KL}[q(\theta) \parallel p(\theta)] \quad (3)$$

where $q(\theta)$ is the variational approximation to the posterior. The first term is the expected negative log-likelihood (NLL) of the data, while the second term is the Kullback–Leibler (KL) regularization, which keeps the posterior close to the prior. The NLL was explicitly expressed for normal predictive distribution of multiple data points (N) as in Equation (4):

$$\mathcal{L}_{\text{NLL}} = -\sum_{i=1}^N \log p(y_i | \mu_i, \sigma_i) = \frac{1}{2} \sum_{i=1}^N \left[\log(2\pi\sigma_i^2) + \frac{(y_i - \mu_i)^2}{\sigma_i^2} \right] \quad (4)$$

where $\mathcal{L}_{\text{NLL}}$ is the loss function corresponding to NLL, which maximizes the likelihood of observed (simulated here) targets (y_i) given the model's predictive mean (μ_i) and variance (σ_i^2) for each data point (i).

All hyperparameters were manually tuned using the training–validation split. The selection of Adam optimizer (learning rate = 0.01) with 1000 epochs and a batch size of 32 yielded the most stable convergence and lowest NLL among tested configurations. Overall, training stability was ensured through feature standardization, KL regularization inherent to the variational layers, and the heteroscedastic likelihood formulation, which jointly prevented divergence and variance collapse during optimization. Early stopping (patience = 40 epochs) was employed as an additional safeguard against overfitting rather than as the sole stabilization mechanism. All computations were conducted using Python 3.11.9, implementing TensorFlow v2.20.0 and TensorFlow Probability v0.25.0, on a system running Windows 11 with a Core i9-13900K CPU and 32 GB of RAM.

Uncertainty Decomposition

During inference, multiple stochastic forward passes were performed, yielding an ensemble of predictions, each with predictive distribution. The predictive mean $\hat{y}(x)$ and total variance $\text{Var}(\hat{y}(x))$ were computed accordingly, as expressed in the following Equations (5) and (6), respectively:

$$\hat{y}(x) = \mathbb{E}_{q(\theta)}[\mu_{\theta}(x)] \quad (5)$$

$$\text{Var}(\hat{y}(x)) = \text{Var}_{q(\theta)}[\mu_{\theta}(x)] + \mathbb{E}_{q(\theta)}[\sigma_{\theta}^2(x)] \quad (6)$$

where $\mu_{\theta}(x)$ and $\sigma_{\theta}^2(x)$ denote the predicted mean and variance, respectively, of the normal likelihood for a given input x using the model of parameters θ . The symbol $(\cdot)$ denotes the model's predicted values to distinguish them from the observed data.

As discussed in Section 1, the first term in Equation (6) represents epistemic uncertainty, which was captured through the stochastic sampling of weights in the DenseVariational layers. The second term represents aleatoric uncertainty, which was modeled by predicting $\sigma_{\theta}(x)$ in the DistributionLambda output layer. Although the calculation of uncertainty is primarily epistemic in conventional BNNs, to capture aleatoric uncertainty, the output layer was defined as a heteroscedastic normal likelihood that jointly predicts the mean and scale parameters of the target (y) distribution. This hybrid structure allowed for the quantification of both uncertainty sources within a single probabilistic model.

Moreover, to ensure the reliability of the predictive variance, we applied a global calibration factor (c) to the predicted standard deviation, after model training was completed. This calibration consists of a single global scaling factor that was estimated using the validation subset only, without modifying the learned BNN parameters or the predictive mean $\hat{y}(x)$. The calibration factor was computed as expressed in Equation (7), rescaling the predicted uncertainty and therefore has limited capacity to overfit.

$$c = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{(y_i - \mu_i)^2}{\sigma_i^2 + \epsilon}} \quad (7)$$

where ϵ is a small constant added for numerical stability. The calibrated variance $(c\sigma)^2$ was subsequently used in all evaluations and uncertainty analyses. Calibration performance was evaluated by comparing proper scoring rules and empirical coverage probabilities on the training, validation, and held-out test subsets (Section 3.2.2) to further assess stability and generalization beyond the data used for calibration.

2.4. Model Application and Design Optimization Process

2.4.1. Sensitivity Analysis Across Climatic Scenarios

We conducted a global sensitivity analysis (SA) using Sobol method to identify the most influential passive design parameters to thermal comfort performance and examine how their relative influence evolves under each climatic scenario. The SA was conducted for the seven climatic scenarios, considering the ten continuous input variables (p1–p10 in Table 3), bounded within the ranges used for database generation (Section 2.2.1) to maintain a feasible design space.

To effectively explore the multidimensional design space, the Sobol quasi-random sampling from SALib Python library was employed with a base size of $N = 2048$, producing $N(2 \times 10 \text{ variables} + 2)$ samples for each climatic scenario (45,056 evaluations). Based entirely on the uncertainty-aware BNN model, the output $\hat{y}(x)$ in Equation (5), which is the predictive mean of 50 stochastic passes, was evaluated for each sampled set of

design parameters. Here, we considered only the mean, since the focus was on building performance sensitivity, not on the SM's own uncertainty. Finally, first-order (S1) and total-order (ST) Sobol indices were computed, where both indices and their 95% confidence intervals were saved for subsequent cross-climatic comparisons.

2.4.2. Resilience-Oriented Building Performance Optimization

As a proof of concept, the optimization was conducted for three selected representative climate scenarios, namely current, RCP4.5–2090 and RCP8.5–2090. In addition, we fixed two orientations of Northwest (NW, the original orientation of the dwelling) and Southeast (SE, representing the opposite orientation). This way, for each optimization run, the nine passive parameters P2–P10 were optimized within their bounds as in Section 2.2.1, to minimize annual pct_hot for a specific climatic scenario and orientation. Since the BNN model produced a predictive distribution of possible pct_hot values for a given design (x), the optimization minimized a robust, risk-aware objective function $J(x)$ (in Equation (8))—integrating predictive uncertainty and model validity—as proxy for the annual pct_hot ($\hat{y}(x)$):

$$J(x) = \text{CVaR}_\alpha [Y(x)] + \lambda_{\text{epi}} \sigma_{\text{epi},\text{norm}}(x) + \gamma \phi(x) \quad (8)$$

The first term, CVaR_α , represents the conditional value at risk of the predicted discomfort distribution $Y(x)$ at a confidence level α of 0.90. This basically measures the predicted discomfort within the worst 10% ($1 - \alpha$ fraction) of the outcomes, thus minimizing tail risk rather than the distribution mean. The distribution of possible outcomes $Y(x)$ was sampled from the SM's predictive normal likelihood distribution $N(\mu_\theta(x), \sigma_\theta^2(x))$ across 128 stochastic passes (different θ realizations) per optimization generation. The second term penalizes the epistemic uncertainty of the SM, expressed as the normalized standard deviation ($\sigma_{\text{epi},\text{norm}}$), to avoid converging around SM under-sampled areas. We normalized each individual epistemic uncertainty ($\sigma_{\text{epi}}(x)$) relative to the median uncertainty of the current population (per generation) as expressed in Equation (9):

$$\sigma_{\text{epi},\text{norm}}(x) = \frac{\sigma_{\text{epi}}(x)}{\text{median}(\sigma_{\text{epi}}) + \varepsilon} \quad (9)$$

where the small constant ε avoids division by zero. The relative contribution of $\sigma_{\text{epi},\text{norm}}(x)$ to the overall objective $J(x)$ was tuned to 0.1, using the weighting coefficient λ_{epi} . This normalization ensures balanced optimization by making the uncertainty penalty scale consistent across all climates, generations, and objective ranges. The third term, $\gamma \phi(x)$, imposes an out-of-distribution (coverage) penalty that prevents the optimizer from exploring regions of underrepresented SM training data. We defined the coverage region in the normalized feature space using a k-Nearest-Neighbor (kNN) distance metric, with $k = 5$.

For each climatic scenario w , we precomputed the coverage baseline of the corresponding normalized training subset ($Z_{\text{train},w}$), by calculating the kNN distance $d_k(z_i)$ for each sample $z_i \in Z_{\text{train},w}$, excluding self matches. We then calculated the 95th percentile to define the coverage radius τ_w ; and the 99th percentile of the residuals ($(d_k(z_i) - \tau_w)_+$) to define the normalization span. Subsequently, during the optimization, the KNN distance $d_k(x)$ to $Z_{\text{train},w}$ for each normalized design candidate (x) was calculated. The coverage penalty $\phi(x)$ was then obtained using a smooth softplus-based ramp, as expressed in Equation (10):

$$\phi(x) = \left[\frac{\log\left(1 + \exp\left[s \frac{d_k(x) - \tau_w}{\text{span}_w}\right]\right)}{\log(1 + \exp(s))} \right]^p \quad (10)$$

where s represents the softplus steepness ($s = 10$), and p is the exponent growth rate, which was set to $p = 2$. Generally, $\phi(x)$ approached zero for design candidates located within the coverage region ($d_k(x) \leq \tau_w$). Moreover, the scaling factor γ was dynamically linked to the population's median CVaR in the first generation, ensuring that the penalty $\phi(x)$ maintained a comparable magnitude to the pct_hot values across all runs. This entailed an increase in $J(x)$ by almost 10–20%, which was fixed across all generations to maintain consistent penalization strength.

The optimization employed the Differential Evolution (DE) algorithm as implemented in *pymoo* Python library using a population of 100 candidates for 100 generations. At each generation, every individual was evaluated through 128 Monte Carlo forward passes, to produce a stable, low-variance estimate of $J(x)$ for parent selection and mutation. The combination of each climate and orientation was optimized independently, where the 20 best solutions of the final generation were initially selected for inverse-scaling and reevaluation in EnergyPlus to verify the surrogate-predicted performance. Subsequently, the final generations from all cases were re-evaluated in EnergyPlus and compared with high-fidelity simulation-based optimizations. This workflow targeted final optimal solutions that can achieve low discomfort values, low epistemic uncertainty, and high in-distribution coverage, underscoring designs that remain robustly comfortable across future climates under SM's uncertainty.

3. Results and Discussion

3.1. Building Energy Simulation

EnergyPlus hourly indoor temperature results for the Bedroom zone were utilized to calculate three adaptive thermal comfort metrics: (1) the percentage of annual hours with overheating conditions (BR_pct_hot) (%); (2) with underheating conditions (BR_pct_cold) (%); and (3) the total discomfort hours (TDH) (h). Table 6 shows an evident progressive intensification of overheating hours in all future scenarios, with a mean BR_pct_hot increasing over a 16-fold (from 1.4% to 22.9% in the extreme RCP 8.5–2090 scenario). In contrast, cold discomfort hours percentage (BR_pct_cold) had almost no impact in the future scenarios, with the maximum value reduced to less than 1% in the 2050 scenarios. Overheating dominates future discomfort, increasing the mean TDH from about 188 h (current) to over 2000 h (RCP8.5–2090).

With regard to both national guidance and international habitability benchmarks, the current overheating percentage (1.40%) falls within the 5% annual deviation from the maximum $T_{o,comf}$, as specified by EN 16798-1 [59] and CIBSE TM52 [60]. Under future climatic scenarios, the sharp increase in mean (22.9%) and maximum (48.25%) overheating duration indicate that operative temperatures would increasingly exceed the 80% acceptability range described by the ASHRAE 55 ACM most of the year. Even though this model allows for higher temperature tolerances of $\sim 31^\circ\text{C}$ for naturally ventilated tropical buildings [49,59], these results suggest that, in the absence of stringent design codes and targeted passive strategies, maintaining acceptable indoor conditions in future Indonesian housing will become increasingly challenging.

Figure 4 shows Pearson's correlations (r) between all design parameters and BR_pct_hot. The strongest negative correlation with thermal discomfort was consistently evident for roof insulation thickness (R_XPS_t), with $r \approx -0.36$ to -0.46 across all scenarios, while positive correlations were found for solar-gain-related parameters, such as WWR, solar absorptance, and SHGC. The magnitude of these correlations increases slightly in future scenarios (compared to the current climate), reflecting an increasing sensitivity to solar gains. In contrast, window U-value (Win_U) exhibited negligible correlations, underscoring the notion that conduction gains are imperceptible relative to solar gains in tropical

regions [34]. Notably, these pairwise correlations reflected only the monotonic relationships between input and output variables. A cross-climatic global sensitivity analysis was still required to quantify the unique contribution of each design parameter to the output variance by accounting for nonlinear interactions and higher-order effects.

Table 6. Simulated thermal comfort metrics in the Bedroom zone in each climatic scenario in Surakarta.

Climate	Output	Mean	STD	Min	25%	Median	75%	Max
Current	BR_pct_hot	1.40	2.59	0.00	0.38	0.68	1.44	29.28
	BR_pct_cold	0.75	0.42	0.08	0.43	0.67	1.00	3.20
	TDH	187.96	234.78	24	85	138	209	2584
RCP 4.5–2050	BR_pct_hot	5.28	4.46	0.58	3.05	4.14	5.92	36.95
	BR_pct_cold	0.13	0.10	0.00	0.06	0.13	0.18	0.92
	TDH	474.01	392.07	52	280	377	537	3238
RCP 4.5–2070	BR_pct_hot	7.29	4.69	0.99	4.40	6.07	8.33	34.45
	BR_pct_cold	0.06	0.05	0.00	0.01	0.05	0.09	0.30
	TDH	643.24	411.40	87	392	540	735	3025
RCP 4.5–2090	BR_pct_hot	9.21	5.43	1.63	5.83	7.81	10.80	37.47
	BR_pct_cold	0.06	0.05	0.00	0.01	0.05	0.08	0.42
	TDH	811.69	476.58	143	515	692	951	3283
RCP 8.5–2050	BR_pct_hot	6.10	4.38	0.86	3.71	5.10	6.92	33.07
	BR_pct_cold	0.10	0.10	0.00	0.03	0.08	0.15	0.89
	TDH	542.95	385.05	75	334	455	616	2905
RCP 8.5–2070	BR_pct_hot	12.46	5.58	3.22	8.93	11.30	14.38	42.44
	BR_pct_cold	0.01	0.02	0.00	0.00	0.00	0.01	0.11
	TDH	1092.2	489.46	282	782	990	1260	3719
RCP 8.5–2090	BR_pct_hot	22.92	5.96	10.54	19.03	22.17	25.89	48.25
	BR_pct_cold	0.00	0.01	0.00	0.00	0.00	0.00	0.06
	TDH	2007.9	522.12	923	1667	1942	2268	4227

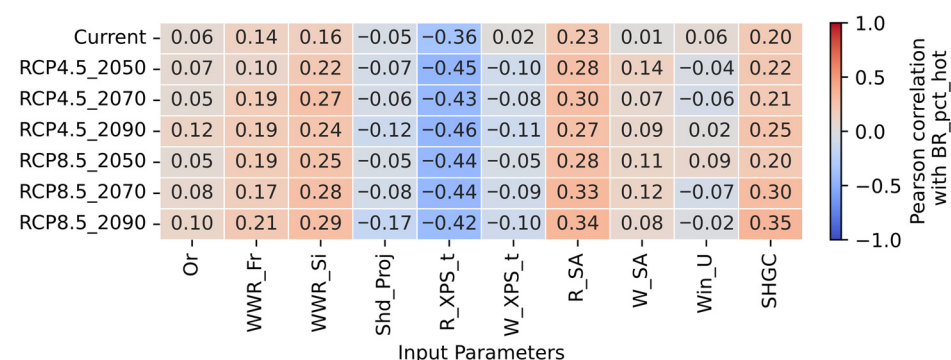


Figure 4. Pearson correlation between passive parameters and thermal comfort outputs.

3.2. Surrogate Model Performance

As mentioned in Section 2.3, we manually tuned the BNN SM's structure and hyperparameters, i.e., the number of DenseVariational layers, neurons per layer, batch size, learning rate, and maximum number of training epochs, to ensure its adequate representation of

the highly non-linear simulated data. Figure 5a illustrates the training history, showing a smooth and stable convergence, in terms of NLL, for both training and validation loss. The smooth evolution, together with the absence of divergence or oscillatory behavior, indicates stable optimization of the BNN model. Note that the training loss leaped to approximately 4.4×10^6 (outside the plotted range) during the first epoch, then instantly stabilized. Additionally, Figure 5b shows a close-up view of the final 80 epochs, revealing an SM convergence around $\text{NLL} \approx 1.2$. The early-stopping criterion ended the training in approximately 78.9 s, after 865 epochs with no validation loss detected as of epoch 825 beyond $\text{NLL} \approx 1.16$.

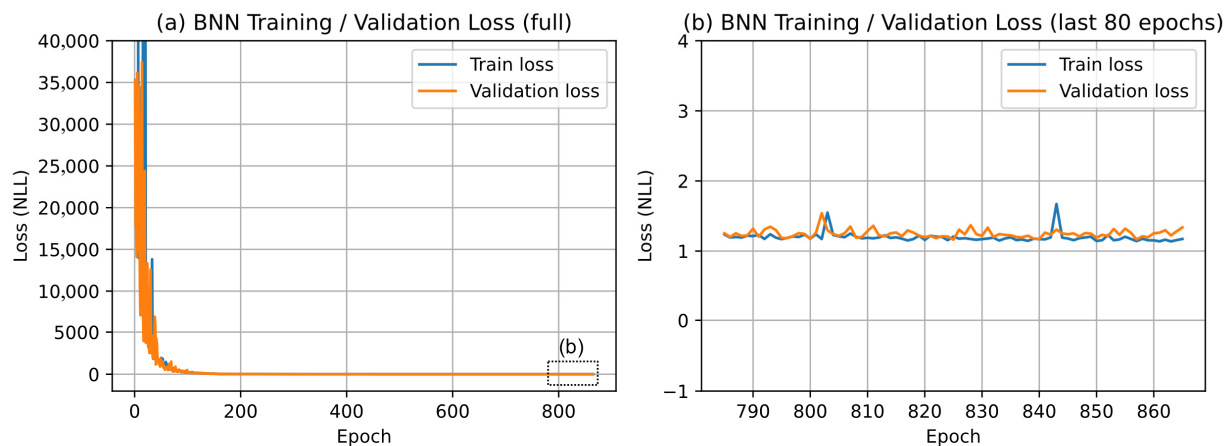


Figure 5. Evolution of the training and validation NLL loss for the BNN model over the full training curve (a) (y-axis is truncated at 4×10^4 for visibility), and the final 80 epochs (b).

3.2.1. Prediction Accuracy

To ensure statistical stability in the BNN model output, we ran 100 stochastic forward passes for each input sample from which the mean and variance of the predictive distributions were computed. Figure 6 illustrates the predictive performance of the BNN on the test subset. Each point represents a case of $\text{BR}_{\text{pct_hot}}$ (mean of predictive distribution $\hat{y}(x)$), while the vertical error bars represent the 95% predictive intervals, derived from the calibrated variance $(\sigma)^2$ (Section Uncertainty Decomposition). The figure confirms a strong predictive fidelity as the points closely aligned to the diagonal line, with a narrow spread of uncertainty bands.

The predictive mean values were used to evaluate deterministic performance metrics (Table 7), where the Coefficient of determination (R^2), RMSE, and MAE were reported for training, validation, and test subsets. The model achieved an $R^2 \approx 0.99$, with RMSE and MAE of approximately 0.8% and 0.5%, respectively, across all subsets. A standard ANN with similar structure and hyperparameters was trained on the same dataset to further attest the comparable accuracy of probabilistic formulation. Since the deterministic ANN model had much fewer trainable parameters, it was tuned separately. Hence, a lower number of neurons per layer and a higher learning rate (of 0.05) were selected to avoid classical overfitting that would occur when using the same settings of the BNN model.

As shown in Table 7, the higher performance of the BNN on the test set confirms that the Bayesian formulation did not degrade the mean performance, despite its larger number of trainable parameters due to the variational layers. Importantly, the close agreement between training, validation, and test performance metrics suggests that the selected model configuration does not suffer from overfitting, despite its high predictive accuracy.

The observed accuracy levels are consistent with those reported in earlier studies employing Bayesian deep learning for building performance surrogates [30,36], notwithstanding differences in climatic conditions, feature representations, and dataset sizes and

types. Importantly, the contribution of the present model extends beyond deterministic accuracy by enabling uncertainty-aware analyses and design applications through fully probabilistic outputs.

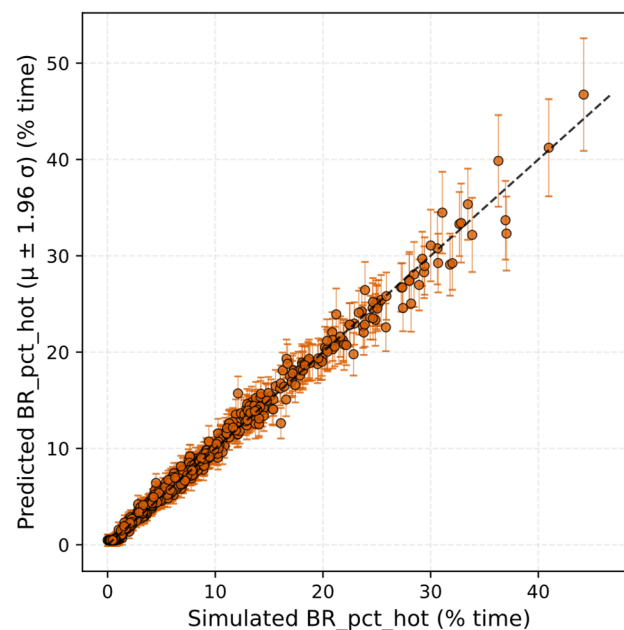


Figure 6. Scatter plot of simulated and SM predicted BR_pct_hot with 95% prediction intervals (on the test subset).

Table 7. Deterministic and calibrated probabilistic performance metrics across all data subsets.

Model	Dataset	R ²	RMSE (%)	MAE (%)	CRPS (%)	PICP ₆₈	PICP ₉₅
Deterministic ANN	Train	0.9925	0.699	0.506	–	–	–
	Valid	0.9893	0.803	0.573	–	–	–
	Test	0.9882	0.884	0.597	–	–	–
BNN	Train	0.9918	0.730	0.473	0.3364	0.7351	0.9727
	Valid	0.9894	0.799	0.512	0.3648	0.7067	0.9429
	Test	0.9905	0.804	0.521	0.3768	0.6876	0.9505

3.2.2. Uncertainty Interpretation

In addition to point-prediction performance, the developed BNN model allowed the decomposition of output variance into its aleatoric (data-intrinsic) and epistemic (model-related) components (Equation (6)). Figure 7 illustrates the relative contribution of both components across the prediction space using the test subset. On average, aleatoric variance contributed to approximately 89% of the total predictive variance (median $\approx 90\%$) across the entire output range. The comparatively small epistemic contribution reflects low sensitivity to parameter uncertainty and indicates stable predictive behavior across weight realizations, which should not be interpreted as excessive model confidence or overfitting. Instead, it primarily results from the dense sampling of the design space used for training. Crucially, this behavior persists when uncertainty is evaluated on the independent test set, confirming that the reduced epistemic variance is not an artifact of training-data memorization. Although these findings confirm that the SM can identify regions—in the design space—of greater predictive uncertainty, the model’s suitability for risk-informed design applications was further verified through structured uncertainty metrics.

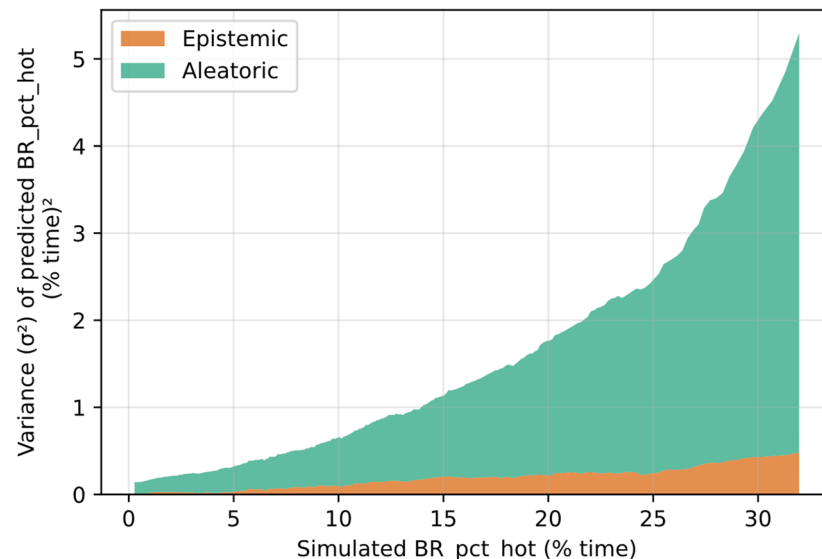


Figure 7. Uncertainty decomposition of aleatoric and epistemic variances along the simulated BR_pct_hot.

In line with the recent recommendations for structured uncertainty evaluation [61], we assessed the reliability of the predictive distributions using the Continuous Ranked Probability Score (CRPS) and the Prediction Interval Coverage Probability (PICP). CRPS resembles the MAE metric but evaluates the entire predictive distribution by comparing the Cumulative Distribution Function (CDF) of the predictions with that of simulations (observed data). PICP, on the other hand, quantifies the empirical coverage rate of the predicted intervals. Hence, PICP_{68} ($\mu \pm 1\sigma$) and PICP_{95} ($\mu \pm 1.96\sigma$) are expected to cover 68% and 95% of the target values, respectively.

Both metrics, summarized in Table 7, were computed for all subsets, achieving CRPS of 0.34–0.38% and empirical coverage probabilities ranging between 0.69 and 0.74 for PICP_{68} , and 0.94–0.97 for PICP_{95} . These values indicate that the predicted distributions are well centered with the simulated data, confirming the BNN model’s reliability for probabilistic inference. Furthermore, the consistency of CRPS and PICP values across training, validation, and test subsets further demonstrates that the uncertainty calibration does not induce subset-specific overfitting and confirms the robust generalization of the SM beyond the data used for calibration. Together, the deterministic and probabilistic results validate the SM’s accuracy and reliability, confirming its suitability for subsequent cross-climatic sensitivity analysis and resilience-oriented design optimization.

3.3. Parameter Sensitivity Shifts Across Scenarios

Global SA of the seven climate scenarios revealed clear patterns of each parameter’s dominance over pct_hot, either by considering their direct effect (S1 index) or interactions with all other parameters (ST index). Table 8 shows S1 and ST values for the key passive parameters across all scenarios, while Figure 8 introduces examples of SA behavior under selected climates, showing the sensitivity shifts over progressively warming projections. Under the current climate, the roof insulation thickness (R_XPS_t) was the most dominant driver of overheating, with an S1 of approximately 0.44, followed by the roof’s solar absorptance (R_SA), with $S1 \approx 0.09$. The combined “direct” influence of these two parameters (~53%) over BR_pct_hot variance, clearly underscored the dominant role of roof’s heat transfer in achieving thermal comfort for such lightweight tropical dwellings. Their total-order indices were even larger, with $ST \approx 0.86$ and 0.44 for R_XPS_t and R_SA, respectively,

representing the dominant sources of interaction-driven variability in thermal comfort performance for this climate scenario.

Table 8. SA results for the five key parameters across all climate scenarios, including the 95% confidence intervals (CI) in parentheses.

Climate Scenario	R_XPS_t		R_SA		SHGC		WWR_Si		Or	
	S1 (CI)	ST (CI)	S1 (CI)	ST (CI)	S1 (CI)	ST (CI)	S1 (CI)	ST (CI)	S1 (CI)	ST (CI)
Current	0.44 (0.13)	0.86 (0.14)	0.09 (0.05)	0.44 (0.11)	0.02 (0.01)	0.05 (0.01)	0.02 (0.02)	0.05 (0.01)	0.02 (0.01)	0.04 (0.01)
RCP4.5_2050	0.47 (0.1)	0.67 (0.08)	0.10 (0.04)	0.29 (0.05)	0.05 (0.02)	0.07 (0.01)	0.07 (0.02)	0.07 (0.01)	0.04 (0.02)	0.05 (0.01)
RCP4.5_2070	0.44 (0.08)	0.6 (0.08)	0.07 (0.03)	0.26 (0.05)	0.06 (0.02)	0.09 (0.01)	0.07 (0.02)	0.09 (0.01)	0.05 (0.02)	0.07 (0.01)
RCP4.5_2090	0.39 (0.06)	0.56 (0.06)	0.10 (0.03)	0.24 (0.04)	0.09 (0.02)	0.11 (0.01)	0.10 (0.02)	0.11 (0.01)	0.06 (0.02)	0.08 (0.01)
RCP8.5_2050	0.45 (0.09)	0.65 (0.09)	0.11 (0.05)	0.28 (0.05)	0.06 (0.02)	0.08 (0.01)	0.07 (0.02)	0.08 (0.01)	0.05 (0.02)	0.06 (0.01)
RCP8.5_2070	0.37 (0.07)	0.48 (0.05)	0.09 (0.02)	0.21 (0.04)	0.11 (0.02)	0.13 (0.01)	0.10 (0.02)	0.11 (0.01)	0.08 (0.02)	0.10 (0.01)
RCP8.5_2090	0.37 (0.06)	0.47 (0.05)	0.09 (0.03)	0.20 (0.03)	0.12 (0.02)	0.12 (0.01)	0.10 (0.02)	0.11 (0.01)	0.09 (0.02)	0.10 (0.01)

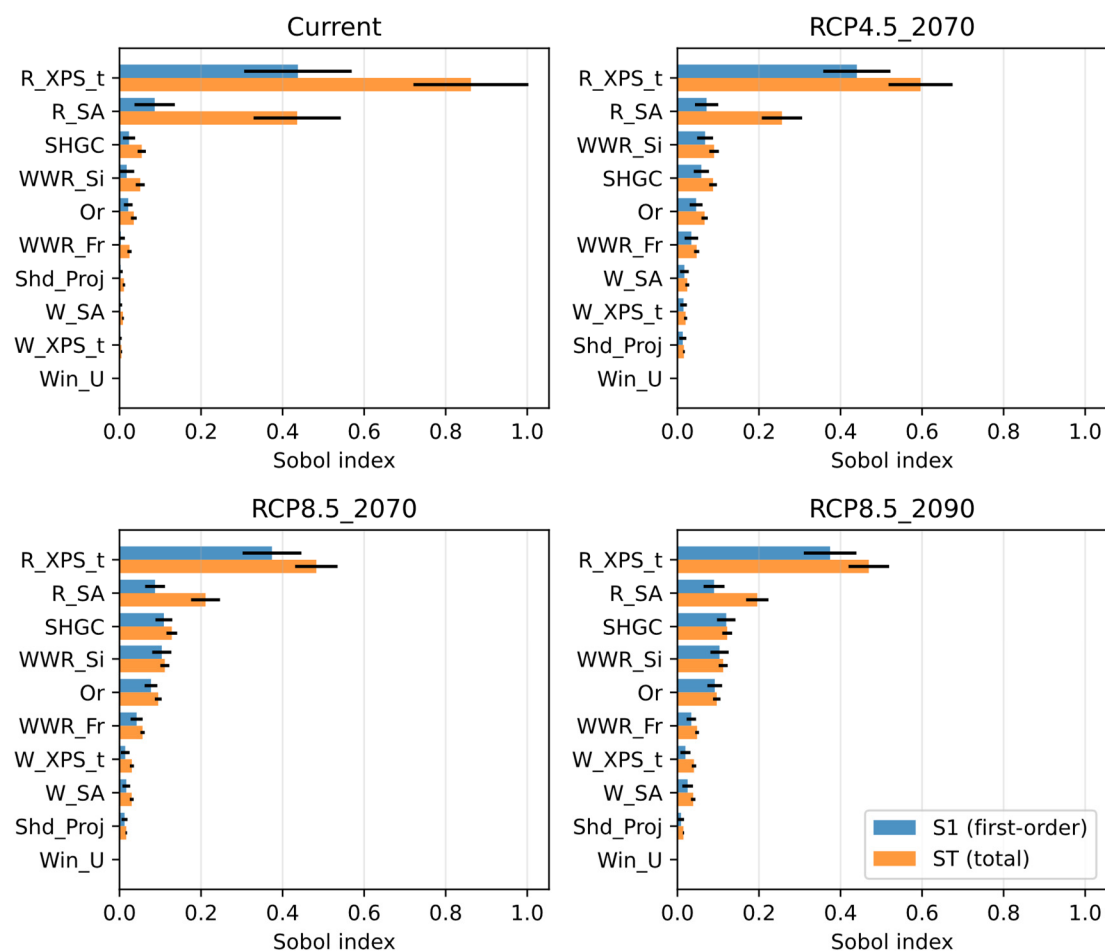


Figure 8. First and Total Sobol indices (S1 and ST) for passive design parameters under four selected scenarios, sorted by their ST. Error bars represent the 95% confidence intervals.

Given the roof's direct exposure to solar radiation and its high thermal conductivity (Table 2), and negligible thermal resistance (metal roofs = zero [62]), conduction heat transfer tends to be the greatest contributor to the indoor overheating, thereby inducing an abrupt increase in *pct_hot*. Both parameters, *R_XPS_t* and *R_SA*, control the thermal resistance of the roof, and hence exhibit the highest importance. Romero et al. [63] utilized Meteonorm to generate six future climatic scenarios, to test the impact of applying EPS insulation to two representative housing typologies in four Mexican cities. They concluded that conduction heat gains through the roof constituted the primary thermal load in both typologies. Hashemi [64] confirmed these findings where iron roofs performed worse than

thatched roofs, regardless of the wall material used. Notwithstanding, Zune et al. [21] found that metal roofs performed as efficiently as thatched roofs only when natural ventilation was available through the false ceiling.

The partial contribution of the façade parameters, WWR and orientation to natural ventilation demonstrates their relatively higher total-order values ($ST < 0.05$) than their first-order effects ($S1 < 0.02$), confirming that their minor contributions were interaction-driven. Importantly, a clear re-distribution of influence among the passive design parameters was evident across the projected climate pathways. Gamero-Salinas et al. [24] found that natural ventilation was the most important parameter affecting overheating in their 3-story representative building in Honduras. However, a prior study [25] confirmed that its potential cooling will decrease in future scenarios. In the present study, as CDD increased from the current scenario ($CDD \approx 1759$ °C·days) to the extreme RCP8.5–2090 scenario ($CDD \approx 3142$ °C·days), outdoor air temperature approaches that of the indoors, where stack pressure is diminished and natural ventilation is thereby suppressed.

Fosas et al. [65] highlighted that higher thermal insulation can be associated with higher overheating duration and severity, given the lack of adequate ventilation. In this case, reducing overheating tended to be driven by intercepting, rather than purging, sensible heat. Consistent with this interpretation, the present results demonstrate the gradual decrease in the “total” influence of roof parameters to nearly half, with ST values of 0.47 and 0.2 for R_{XPS_t} and R_{SA} , respectively. In contrast, the parameters associated with solar gains and façade geometry, particularly, SHGC, side façade WWR (WWR_{Si}), and orientation (Or), exhibited a notable consistent increase in influence, with S1 values surpassing that of R_{SA} in RCP8.5–2070 and 2090.

This SA multi-scenario cross-evaluation demonstrated that controlling the conductive and radiative heat exchange through the roof alone remained effective for mitigating thermal discomfort in the present scenario. However, the notable shift in thermal performance sensitivity indicated that thermal resilience was more dependent on façade-driven solar control measures in future scenarios. Furthermore, the gap between S1 and ST, tended to diminish in warmer scenarios (Table 8). This rather increasing direct influence of most parameters suggests that the optimization of their individual influence could be as effective in future design guidelines for tropical housing as optimizing their synergetic influence. In the cross-climatic BPO stage, however, we retained all the parameters, to avoid excluding any potential secondary interactions.

3.4. Surrogate Uncertainty-Aware Optimal Solutions

Across all scenarios, the uncertainty-aware objective $J(x)$ consistently guided the optimization toward designs with low BR_pct_hot , while exploring design-space regions close to the training data distribution ($\phi(x) \approx \text{zero}$). For example, as shown in Table 9, the most extreme scenario (RCP8.5–2090, NW-oriented bedroom) yielded optimal designs with a predictive μ of approximately 10%, with relatively tight predictive distributions ($CVaR_{90\%} \approx 11.95\%$, $\sigma \approx 0.9\%$). Furthermore, the optimal design cases in less extreme scenarios exhibited almost negligible overheating time (predicted μ between 0.47% and 1.9%), in agreement with previous findings that passive measures remain highly effective under current –or mildly warmer– climatic conditions [24,63]. Supplementary Material S3 provides a summary of the surrogate-assisted optimal solutions in six optimization scenarios, as sorted by $J(x)$.

Table 9. The most robust five surrogate-assisted optimal solutions for the case RCP8.5–2090, NW. μ and σ denote the BNN’s predictive mean and standard deviation, respectively, while “sim” represents the re-evaluated BR_pct_hot using EnergyPlus.

#	Design Parameters										J(x)	CVaR	ϕ	μ	σ	sim
	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10						
1	343.4	0.15	0.2	1.49	0.06	0.15	0.1	0.12	0.35	0.11	12.22	11.93	0	10.05	0.89	10.26
2	343.4	0.15	0.2	1.41	0.06	0.14	0.1	0.11	0.34	0.11	12.23	11.94	0	10.08	0.88	10.23
3	343.4	0.15	0.2	1.43	0.06	0.14	0.1	0.11	0.20	0.11	12.24	11.95	0	10.03	0.89	10.24
4	343.4	0.15	0.2	1.49	0.04	0.15	0.1	0.12	0.17	0.11	12.25	11.95	0	9.98	0.91	10.17
5	343.4	0.15	0.2	1.46	0.05	0.14	0.1	0.10	0.56	0.11	12.25	11.96	0	10.09	0.88	10.22

Note: Shaded columns indicate the variables used for direct comparison in Section 3.4.1.

3.4.1. Validation of Surrogate-Derived Optimal Solutions

Previous studies verified surrogate-assisted optimal solutions, given that SMs may produce misleading outputs in under-represented regions of the design space [66,67]. EnergyPlus re-evaluation of the final generation is therefore recommended to provide a more reliable basis for assessing the robustness of the SM-guided search and identifying potential false optima. To contextualize the benefits of the uncertainty-aware surrogate optimization, we employed both the deterministic ANN model (Section 3.2.1) and the BNN model in six validation scenarios. Figure 9 exemplifies this validation, by comparing the ANN’s outputs and BNN’s predictive μ against the (re)simulated BR_pct_hot, for the RCP8.5–2090 scenario. Overall, the BNN predictions for the optimal design cases were congruent with their simulated counterparts, with the latter consistently falling within the BNN’s 95% predictive intervals ($\sigma \approx 0.8\%$). the ANN model, on the other hand, substantially underestimated overheating hours in both orientations.

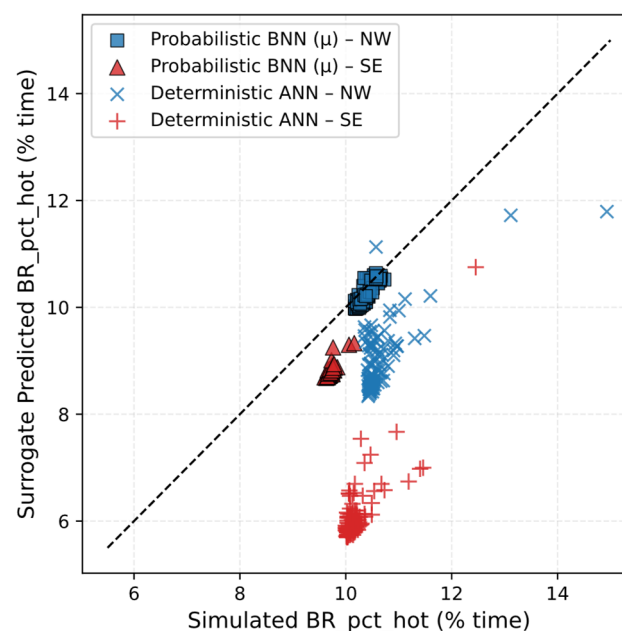


Figure 9. Scatter plot showing the DE final generation of surrogate predictions (BNN μ and ANN) and their re-evaluations, under RCP8.5–2090 scenario, for NW (blue) and SE (red) orientations.

Across the six optimization scenarios (600 DE individuals), the MAE was around 0.33% with 90th percentile absolute error (AE₉₀) of ~0.9% and RMSE of ~0.43% (Table 10). The top surrogate-optimal solutions likewise exhibited absolute differences within 0.32%, with only a few individuals deviating by approximately 1%, which is typically associated with higher

$\phi(x)$. Contrarily, the ANN model showed higher MAE and RMSE than those of the BNN model, by a factor of five. Overall, the re-evaluated performance of the BNN model did not show any over-confidence in optimal regions—unlike deterministic surrogate models—thus supporting the use of the uncertainty-aware objective function, which suppresses overconfident predictions and eliminates irrational optima.

Table 10. Aggregate predictive error of the BNN and ANN surrogates, computed over 600 individuals from the DE final generations.

Model/Error	MAE	RMSE	AE ₉₀
BNN	0.3309	0.4321	0.901
ANN	1.7812	2.2289	4.1615

3.4.2. Comparison with Simulation-Based Optimization

In this stage, EnergyPlus-based optimizations served as benchmarks for the quality of the surrogate-assisted re-evaluated solutions. The optimizations were performed in Grasshopper for 50 generations, each with a population of 50, employing the self-adaptive DE algorithm with ensemble of mutation Strategies (jEDE). Across the six validation scenarios, the surrogate-assisted re-evaluated designs were highly congruent with those obtained in EnergyPlus-based optimization, with differences in BR_pct_hot ranging between 0.01% and 0.14% (Table 11).

Table 11. Minimum BR_pct_hot comparison between surrogate-assisted optimization (after re-evaluation) and simulation-based (EnergyPlus) optimization approaches for all validation scenarios.

Climate Scenario	Orientation	Surrogate-Assisted (BNN) Re-Evaluated BR_pct_hot	Simulation-Based (EnergyPlus) BR_pct_hot
Current	NW	0.09	0.02
	SE	0.03	0.02
RCP4.5–2090	NW	2.04	1.99
	SE	1.83	1.75
RCP8.5–2090	NW	10.00	9.93
	SE	9.60	9.46

Beyond the objective space agreement, the parallel coordinate plots in Figure 10 show the top twenty design cases from both optimization approaches, with the design parameters sorted according to their ST values (Section 3.3). The design parameters of both approaches consistently prioritized designs with low WWRs, SHGC, and solar absorptance, in the RCP8.5–2090 scenario (Figure 10c). The same solutions exhibited low-to-moderate roof insulation thickness (R_XPS_t), and high wall insulation (W_XPS_t) and shading projection (Shd_proj). These findings are in agreement with [25], regarding the WWRs and with [68], regarding shading projection, windows U-value and roof and wall insulations. In contrast, in moderately warm climatic conditions (Figure 10b), the parameters with lower sensitivity values, such as Shd_proj, W_XPS_t, and Win_U, exhibited wider parameter ranges, indicating that optimum pct_hot could be achieved by multiple parameter combinations. This effect was even more pronounced in the current climate (Figure 10a), where a broader set of design configurations can deliver nearly identical pct_hot values, approaching the simulation-based optima.

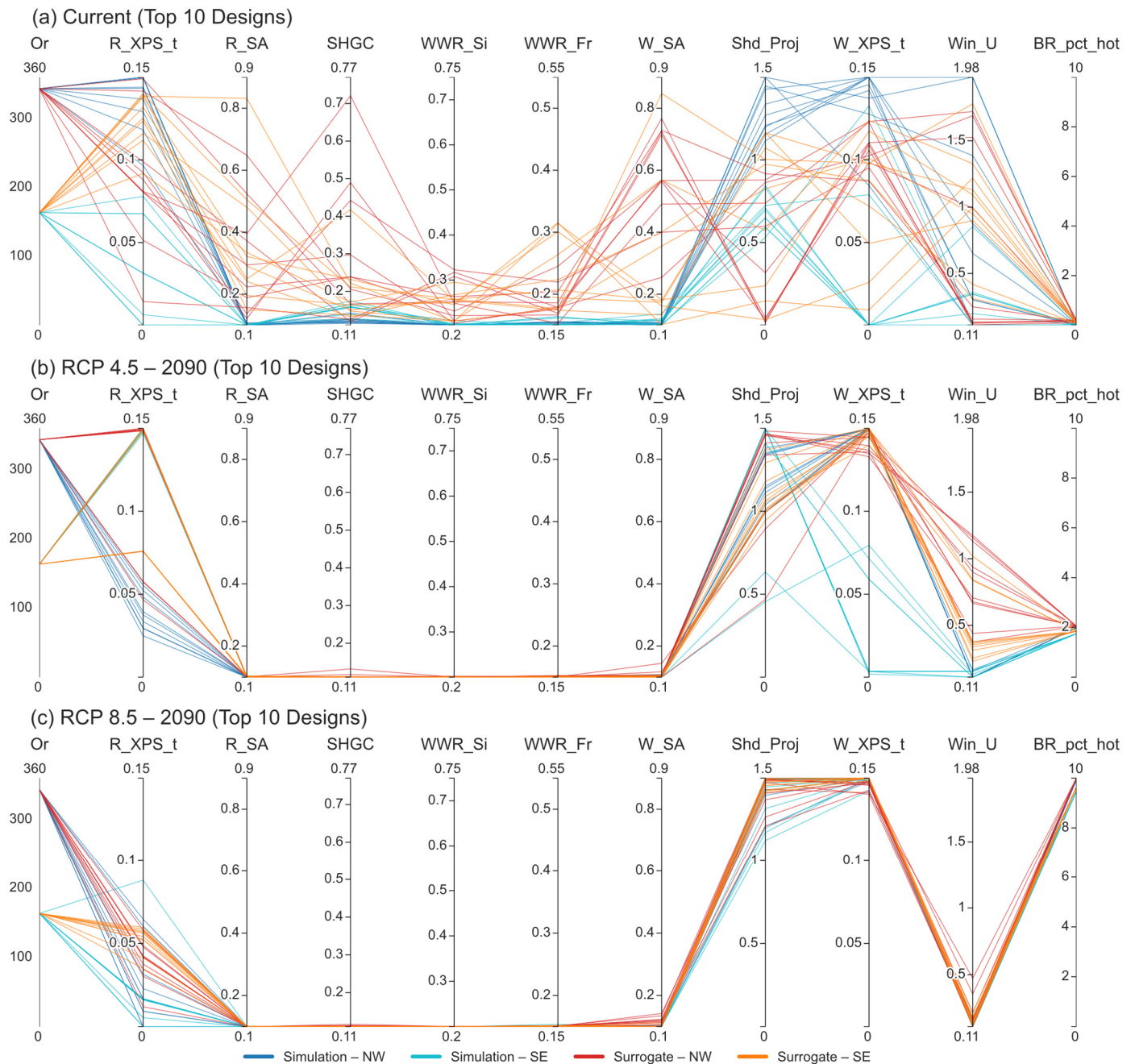


Figure 10. Parallel coordinate plots illustrating the top ten design solutions obtained from simulation-based and surrogate-assisted optimization for the (a) current, (b) RCP4.5–2090, and (c) RCP8.5–2090 scenarios.

These results, together with the sensitivity analysis (Section 3.3), highlight potential limitations in current national design guidelines when evaluated under progressively warming climates. For example, the Indonesian National Standard SNI-6389-2020 [69] defines fixed threshold values of 35 W/m^2 for Roof and Overall Thermal Transfer Values (RTTV and OTTV), implicitly assuming balanced and static contributions from the three heat gain components: wall/roof conduction, window conduction, and solar heat gain through fenestrations, neglecting the varying influence of the design strategies. Moreover, in case of failure to fulfill the RTTV/OTTV requirements, the standard defines a set of descriptive guidelines, for instance, to reduce solar absorptance, WWR, SHGC and use external shade. Although the present findings align with the guidelines, the relative importance of these strategies was observed to shift substantially across future climate

scenarios. This suggests that fixed-weight performance indicators may become increasingly misaligned with actual overheating risk under climate change.

It is important to emphasize that the quantitative results presented here are specific to the studied dwelling and climatic context and are not directly generalizable to other building typologies or regions. Instead, the case study demonstrates a transferable surrogate-assisted, uncertainty-aware optimization approach. The proposed framework can be applied to other case studies to systematically scrutinize the assumptions embedded in existing building codes and guidelines and to explore climate-responsive prioritization of passive design strategies in the future, while maintaining a computationally scalable manner.

Overall, the adopted surrogate-assisted optimization entailed similar design tendencies to those obtained by the simulation-based approach, with a wide range of design alternatives, particularly those with the least influencing parameters. This approves the uncertainty-aware optimization objective—introduced by the probabilistic BNN—as a feasible and overconfidence-resistant optimization method of passive design strategies. Despite being the highest-fidelity benchmark, the elapsed time of a single simulation-based optimization run, in this study, approached 3 h and 10 min, proving computationally prohibitive, especially for multi-scenario investigations. Contrarily, surrogate-assisted optimizations of 100 DE generations required almost 7.5 min, with additional 4 min for re-evaluating the final generation—a 96% reduction in computation time. It is worth noting that the computational cost of the uncertainty-aware surrogate-assisted workflow is higher than typical deterministic surrogate optimization processes. This stems from the multiple stochastic forward passes required to estimate epistemic and aleatoric uncertainty and compute the tail-risk objective. Although this overhead could be significantly diminished by using fewer stochastic passes (e.g., below 128), doing so would compromise BNN prediction robustness, as limited sampling weakens uncertainty quantification and increases overconfidence in under-represented regions of the design space. Hence, the adopted configuration reflects a deliberate tradeoff between computational efficiency and reliable uncertainty-aware optimization.

4. Conclusions

This study demonstrated that uncertainty-aware surrogate-assisted optimization can substantially reduce computational demand while improving reliability, offering a practical tool to support early-stage decisions in climate-resilient housing design. By incorporating predictive uncertainty through Bayesian Neural Networks (BNNs), the framework not only achieves high predictive accuracy but also mitigates overconfident surrogate behavior that may otherwise mislead optimization toward false optima. This robust performance of the BNN across current and future climates indicates that uncertainty-aware probabilistic surrogates are better suited for multi-climatic generalization than deterministic Artificial Neural Networks (ANNs), whose reliability tends to deteriorate in high-dimensional design spaces and under future climatic scenarios.

The strong agreement between surrogate-assisted optima and optimization-based reevaluations confirms that probabilistic surrogates can replace large numbers of computationally expensive simulations without compromising solution quality. Combined with the uncertainty-aware sensitivity analysis, the framework enables climate-informed parameter ranking, thus supporting more robust design decisions.

The analysis revealed that thermal resilience in affordable tropical housing is primarily governed by roof insulation thickness, roof and wall solar absorptance, window-to-wall ratio, and glazing solar heat gain coefficient. While roof-related parameters dominate overheating mitigation under current conditions, future scenarios show a redistribution of influ-

ence toward solar-gain-related variables and façade geometry. As overheating intensifies, the reduction in design resilience margins—relative to the broader range typically observed under current conditions—signifies the need for adaptive, climate-responsive prioritization of passive design strategies, rather than reliance on purely prescriptive guidelines.

The proposed uncertainty-aware workflow provides a structured method for integrating physics-based simulation, probabilistic surrogate modeling, and multi-scenario optimization within a unified decision-support framework. While the numerical findings of the present study are specific to the selected case dwelling, the primary contribution lies in the transferable methodology rather than direct generalization of results. The framework is directly compatible with parametric design environments and can be extended to other building typologies, locations, and regulatory contexts following appropriate model setup and calibration. By enabling efficient exploration of large design spaces under multiple future climate scenarios while explicitly accounting for surrogate uncertainty, the proposed framework offers a scalable foundation for bridging detailed simulation studies and policy-relevant assessments, supporting evidence-based justification of existing design guidelines.

Ultimately, the findings indicate that probabilistic SM optimization tools can support the development of resilience-oriented design strategies for tropical housing. As climate warming exacerbates thermal risks, methods that quantify and manage predictive uncertainty will be of vital importance for guiding design decisions and promoting the practice of performance-based environmental design.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/buildings16020328/s1>. S1: Energy Management System (EMS) and Operability Assumptions; S2: Generation of Future Climate Scenarios [70–72]; S3: Top 20 surrogate-assisted optimal solutions per climate–orientation case, ranked by $J(x)$.

Author Contributions: Conceptualization, I.E. and Y.I.; methodology, I.E. and Y.I.; software, I.E.; validation, I.E. and X.Z.; investigation, I.E., X.Z. and F.Z.M.; writing—original draft preparation, I.E., X.Z. and F.Z.M.; writing—review and editing, Y.I. and A.H.; visualization, I.E.; supervision, A.H.; project administration, A.H.; funding acquisition, A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Kyushu University Platform of Inter-/Transdisciplinary Energy Research (Q-PIT) through its “Module-Research Program”.

Data Availability Statement: The simulated datasets used for training the BNN model are openly available in Zenodo at DOI: 10.5281/zenodo.18136581 [54].

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sari, D.P. A Review of How Building Mitigates the Urban Heat Island in Indonesia and Tropical Cities. *Earth* **2021**, *2*, 653–666. [CrossRef]
2. UN-HABITAT. *World Cities Report 2022: Envisaging the Future of Cities*; United Nations: New York, NY, USA, 2022.
3. Uchôa, J.G.S.M.; Bertotto, L.E.; dos Santos, M.S.; Reis, A.; Mendiondo, E.M.; Wendland, E.C. Tropical cities research boundaries: A bibliometric analysis to bridge the gaps through multi-dimensional and cross-disciplinary features. *Front. Sustain. Cities* **2023**, *5*, 1154667. [CrossRef]
4. UN-HABITAT. *World Cities Report 2024: Cities and Climate Action*; United Nations: New York, NY, USA, 2024.
5. Mazáček, D. Affordable Housing: Concept and Affordability Measurements. *Socio-Econ. Humanit. Stud.* **2024**, *19*, 5–30. [CrossRef]
6. Kokatnur, T.; Ferreira, S.; Akkoç, B.K.; Markarian, E.; Nojedehe, P.; Qiblawi, S.; Sewraj, K.; Gunay, B.; O’Brien, W.; Papineau, M.; et al. A review of passive design strategies and their effect on thermal resilience in low-income households. *Energy Build.* **2025**, *348*, 116508. [CrossRef]
7. Kajjoba, D.; Wesonga, R.; Lwanyaga, J.D.; Kasedde, H.; Olupot, P.W.; Kirabira, J.B. Assessment of thermal comfort and its potential for energy efficiency in low-income tropical buildings: A review. *Sustain. Energy Res.* **2025**, *12*, 25. [CrossRef]

8. Kajjoba, D.; Kasedde, H.; Olupot, P.W.; Lwanyaga, J.D. Evaluation of thermal comfort and air quality of low-income housing in Kampala City, Uganda. *Energy Build. Environ.* **2022**, *3*, 508–524. [\[CrossRef\]](#)
9. Simões, G.M.F.; Leder, S.M.; Labaki, L.C. How uncomfortable and unhealthy can social (low-cost) housing in Brazil become with use? *Build. Environ.* **2021**, *205*, 108218. [\[CrossRef\]](#)
10. Dodman, D.; Hayward, B.; Pelling, M.; Castan Broto, V.; Chow, W.; Chu, E.; Dawson, R.; Khirfan, L.; McPhearson, T.; Prakash, A.; et al. Cities, Settlements and Key Infrastructure. In *IPCC, 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Pörtner, H.O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., et al., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2022; pp. 907–1040.
11. Jayalath, A.; Vaz-Serra, P.; Hui, F.K.P.; Aye, L. Thermally comfortable energy efficient affordable houses: A review. *Build. Environ.* **2024**, *256*, 111495. [\[CrossRef\]](#)
12. Wei, J.; Li, H.X.; Sadick, A.-M.; Noguchi, M. A systematic review of key issues influencing the environmental performance of social housing. *Energy Build.* **2024**, *319*, 114566. [\[CrossRef\]](#)
13. Murtyas, S.; Hagishima, A.; Kusumaningdyah, N. On-site measurement and evaluations of indoor thermal environment in low-cost dwellings of urban Kampung district. *Build. Environ.* **2020**, *184*, 107239. [\[CrossRef\]](#)
14. Bhikhoo, N.; Hashemi, A.; Cruickshank, H. Improving Thermal Comfort of Low-Income Housing in Thailand through Passive Design Strategies. *Sustainability* **2017**, *9*, 1440. [\[CrossRef\]](#)
15. Hashemi, A. *Effects of Natural Ventilation on Thermal Comfort in Low-Income Tropical Housing*; Zero Energy Mass Custom Home (ZEMCH) 2019 International Conference Proceedings: Seoul, Republic of Korea, 2019; pp. 237–244.
16. Hernández, G.; Cetina-Quñones, A.; Bassam, A.; Carrillo, J. Passive strategies towards energy efficient social housing: A parametric case study and decision-making framework in the Mexican tropical climate. *J. Build. Eng.* **2024**, *82*, 108282. [\[CrossRef\]](#)
17. Lapisa, R.; Arwizet; Kurniawan, A.; Krismadinata; Rahman, H.; Romani, Z. Optimized Design of Residential Building Envelope in Tropical Climate Region: Thermal Comfort and Cost Efficiency in an Indonesian Case Study. *J. Arch. Eng.* **2022**, *28*, 05022002. [\[CrossRef\]](#)
18. Handayani, K.N.; Murtyas, S.; Wijayanta, A.T.; Hagishima, A. Thermal Comfort Challenges in Home-Based Enterprises: A Field Study from Surakarta's Urban Low-Cost Housing in a Tropical Climate. *Sustainability* **2024**, *16*, 6838. [\[CrossRef\]](#)
19. Hashemi, A. Effects of thermal insulation on thermal comfort in low-income tropical housing. *Energy Procedia* **2017**, *134*, 815–824. [\[CrossRef\]](#)
20. Hashemi, A.; Khatami, N. Effects of solar shading on thermal comfort in low-income tropical housing. *Energy Procedia* **2017**, *111*, 235–244. [\[CrossRef\]](#)
21. Zune, M.; Rodrigues, L.; Gillott, M. Vernacular passive design in Myanmar housing for thermal comfort. *Sustain. Cities Soc.* **2020**, *54*, 101992. [\[CrossRef\]](#)
22. Tantasavasdi, C.; Chenvidyakarn, T.; Pichaisak, M. Integrative passive design for climate change: A new approach for tropical house design in the 21st century. *Int. J. Build. Urban Inter. Landsc. Technol.* **2011**, *1*, 5–20. [\[CrossRef\]](#)
23. Callejas, I.J.A.; Apolonio, R.M.; da Guarda, E.L.A.; Durante, L.C.; de Andrade Carvalho Rossetti, K.; Roseta, F.; Amarante, L.M.D. Bermed Earth-Sheltered Wall for Low-Income House: Thermal and Energy Measure to Face Climate Change in Tropical Region. *Appl. Sci.* **2021**, *11*, 420.
24. Gamero-Salinas, J.; Monge-Barrio, A.; Kishnani, N.; López-Fidalgo, J.; Sánchez-Ostiz, A. Passive cooling design strategies as adaptation measures for lowering the indoor overheating risk in tropical climates. *Energy Build.* **2021**, *252*, 111417. [\[CrossRef\]](#)
25. Liu, S.; Kwok, Y.T.; Lau, K.K.-L.; Ouyang, W.; Ng, E. Effectiveness of passive design strategies in responding to future climate change for residential buildings in hot and humid Hong Kong. *Energy Build.* **2020**, *228*, 110469. [\[CrossRef\]](#)
26. Murtyas, S.; Qian, R.; Matsuo, T.; Tuck, N.W.; Zaki, S.A.; Hagishima, A. Thermal comfort in a two-storey Malaysian terrace house: Are passive cooling methods sufficient in present and future climates? *J. Build. Eng.* **2024**, *96*, 110412. [\[CrossRef\]](#)
27. Ali, A.; Jayaraman, R.; Mayyas, A.; Alaifan, B.; Azar, E. Machine learning as a surrogate to building performance simulation: Predicting energy consumption under different operational settings. *Energy Build.* **2023**, *286*, 112940. [\[CrossRef\]](#)
28. Elwy, I.; Hagishima, A. The artificial intelligence reformation of sustainable building design approach: A systematic review on building design optimization methods using surrogate models. *Energy Build.* **2024**, *323*, 114769. [\[CrossRef\]](#)
29. Westermann, P.; Evins, R. Surrogate modelling for sustainable building design—A review. *Energy Build.* **2019**, *198*, 170–186. [\[CrossRef\]](#)
30. Westermann, P.; Evins, R. Using Bayesian deep learning approaches for uncertainty-aware building energy surrogate models. *Energy AI* **2021**, *3*, 100039. [\[CrossRef\]](#)
31. van der Meer, D.W.; Shepero, M.; Svensson, A.; Widén, J.; Munkhammar, J. Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Processes. *Appl. Energy* **2018**, *213*, 195–207. [\[CrossRef\]](#)

32. Gengembre, E.; Ladevie, B.; Fudym, O.; Thuillier, A. A Kriging constrained efficient global optimization approach applied to low-energy building design problems. *Inverse Probl. Sci. Eng.* **2012**, *20*, 1101–1114. [\[CrossRef\]](#)
33. Chen, R.; Tsay, Y.-S.; Zhang, T. A multi-objective optimization strategy for building carbon emission from the whole life cycle perspective. *Energy* **2023**, *262*, 125373. [\[CrossRef\]](#)
34. Elwy, I.; Hagishima, A. Surrogate-assisted Passive Design Exploration of Residential Thermal Comfort in Multiple Climates. In Proceedings of the ASim Conference 2024: 5th Asia Conference of IBPSA, Osaka, Japan, 8–10 December 2024; Volume 5, pp. 467–474.
35. Bamdad, K.; Cholette, M.E.; Bell, J. Building energy optimization using surrogate model and active sampling. *J. Build. Perform. Simul.* **2020**, *13*, 760–776. [\[CrossRef\]](#)
36. Bampoulas, A.; Pallonetto, F.; Mangina, E.; Finn, D.P. A Bayesian deep-learning framework for assessing the energy flexibility of residential buildings with multicomponent energy systems. *Appl. Energy* **2023**, *348*, 121576. [\[CrossRef\]](#)
37. Chong, A.; Menberg, K. Guidelines for the Bayesian calibration of building energy models. *Energy Build.* **2018**, *174*, 527–547. [\[CrossRef\]](#)
38. Es, T.; Yi, Z.; Alexander, I.J.F. Acceleration of Building Design Optimisation Through the Use of Kriging Surrogate Models. In Proceedings of the BSO Conference 2012: 1st Conference of IBPSA-England, Loughborough, UK, 10–11 September 2012; Volume 1, pp. 1–8.
39. Gilan, S.S.; Goyal, N.; Dilkina, B. Active Learning in Multi-objective Evolutionary Algorithms for Sustainable Building Design. In Proceedings of the Genetic and Evolutionary Computation Conference 2016, Denver, CO, USA, 20–24 July 2016; pp. 589–596.
40. Lahmar, S.; Maalmi, M.; Idchabani, R. Multiobjective building design optimization using an efficient adaptive Kriging metamodel. *Simulation* **2023**, *101*, 557–573. [\[CrossRef\]](#)
41. Minami, Y.; Kusumaningdyah, N.H.; Komariah, K.; Murtyas, S.; Hagishima, A. In Field Observation of Indoor Environment in Housing from Slum-Upgrading Projects in Surakarta City, Indonesia. In Proceedings of the 14th International Symposium on City Planning and Environmental Management in Asian Countries, Fukuoka, Japan, 11–13 January 2025; Asian Urban Research Group (AURG): Fukuoka, Japan, 2025.
42. Febrinastri, F.; Fadilah, R. *Pre-Printed Technology for Earthquake-Resistant Simple Homes*; Arkadia Digital Media: Jakarta, Indonesia, 2021. (In Indonesian)
43. Muhammed, F.Z.; Yamaguchi, K.; Handayani, K.N.; Hagishima, A. Affordable Housing in Developing Regions: A Systematic Review of Materials, Methods and Critical Success Factors with Case Insights. *Buildings* **2025**, *15*, 4015. [\[CrossRef\]](#)
44. Roudsari, M.S.; Mackey, C. Ladybug-Tools. Available online: <https://www.ladybug.tools/> (accessed on 10 January 2025).
45. Coakley, D.; Raftery, P.; Keane, M. A review of methods to match building energy simulation models to measured data. *Renew. Sustain. Energy Rev.* **2014**, *37*, 123–141. [\[CrossRef\]](#)
46. METEONORM, Irradiation Data for Every Place on Earth. Available online: <https://meteonorm.com/> (accessed on 11 October 2025).
47. Feriadi, H.; Wong, N.H. Thermal comfort for naturally ventilated houses in Indonesia. *Energy Build.* **2004**, *36*, 614–626. [\[CrossRef\]](#)
48. Damiati, S.A.; Zaki, S.A.; Rijal, H.B.; Wonorahardjo, S. Field study on adaptive thermal comfort in office buildings in Malaysia, Indonesia, Singapore, and Japan during hot and humid season. *Build. Environ.* **2016**, *109*, 208–223. [\[CrossRef\]](#)
49. Parkinson, T.; de Dear, R.; Brager, G. Nudging the adaptive thermal comfort model. *Energy Build.* **2020**, *206*, 109559. [\[CrossRef\]](#)
50. Nguyen, A.T.; Singh, M.K.; Reiter, S. An adaptive thermal comfort model for hot humid South-East Asia. *Build. Environ.* **2012**, *56*, 291–300. [\[CrossRef\]](#)
51. de Dear, R.J.; Brager, G.S. Thermal comfort in naturally ventilated buildings: Revisions to ASHRAE Standard 55. *Energy Build.* **2002**, *34*, 549–561. [\[CrossRef\]](#)
52. Lamsal, P.; Bajracharya, S.B.; Rijal, H.B. A Review on Adaptive Thermal Comfort of Office Building for Energy-Saving Building Design. *Energies* **2023**, *16*, 1524. [\[CrossRef\]](#)
53. Yao, R.; Zhang, S.; Du, C.; Schweiker, M.; Hodder, S.; Olesen, B.W.; Toftum, J.; Romana d’Ambrosio, F.; Gebhardt, H.; Zhou, S.; et al. Evolution and performance analysis of adaptive thermal comfort models—A comprehensive literature review. *Build Environ.* **2022**, *217*, 109020. [\[CrossRef\]](#)
54. Elwy, I. Simulated Dataset for Thermal Resilience Assessment of Affordable Housing in Surakarta, Indonesia under Current and Future Climate Scenarios [Data Set]. Zenodo. 2026. Available online: <https://zenodo.org/records/18136581> (accessed on 11 October 2025).
55. Westermann, P.; Welzel, M.; Evins, R. Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones. *Appl. Energy* **2020**, *278*, 115563. [\[CrossRef\]](#)
56. Elwy, I.; Hagishima, A. Optimizing energy predictions: A generalized surrogate modeling approach for diverse building designs and climates. In Proceedings of the Building Simulation 2025: 18th Conference of IBPSA, Brisbane, Australia, 24–27 August 2025; Volume 19.

57. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
58. Dillon, J.V.; Langmore, I.; Tran, D.; Brevdo, E.; Vasudevan, S.; Moore, D.; Patton, B.; Alemi, A.; Hoffman, M.; Saurous, R.A. TensorFlow Distributions. *arXiv* **2017**, arXiv:1711.10604. [[CrossRef](#)]
59. EN 16798-1; Indoor Environmental Input Parameters for Design and Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics. British Standards Institution, BSI: London, UK, 2019.
60. Chartered Institution of Building Services Engineers (CIBSE). *The Limits of Thermal Comfort: Avoiding Overheating*; CIBSE: London, UK, 2013.
61. Xu, X.; Hu, Y.; Atamturktur, S.; Chen, L.; Wang, J. Systematic review on uncertainty quantification in machine learning-based building energy modeling. *Renew. Sustain. Energy Rev.* **2025**, *218*, 115817. [[CrossRef](#)]
62. ASHRAE. *ASHRAE Handbook of Fundamentals (SI Edition)*; American Society of Heating, Refrigerating, and Air-Conditioning Engineers: Atlanta, Georgia, 2017.
63. Romero, D.; Torres, K.A.; Gonzalez, J.; Cetina-Quiñones, A.J.; Acosta, C.; Sadoqi, M.; Bassam, A. Climate Change and Assessing Thermal Comfort in Social Housing of Southeastern Mexico: A Prospective Study Using Machine Learning and Global Sensitivity Analysis. *Sustainability* **2025**, *17*, 9596. [[CrossRef](#)]
64. Hashemi, A. Climate Resilient Low-Income Tropical Housing. *Energies* **2016**, *9*, 468. [[CrossRef](#)]
65. Fosas, D.; Coley, D.A.; Natarajan, S.; Herrera, M.; de Pando, M.F.; Ramallo-Gonzalez, A. Mitigation versus adaptation: Does insulating dwellings increase overheating risk? *Build. Environ.* **2018**, *143*, 740–759. [[CrossRef](#)]
66. Kubwimana, B.; Najafi, H. A novel approach for optimizing building energy models using machine learning algorithms. *Energies* **2023**, *16*, 1033. [[CrossRef](#)]
67. Shen, C.; Lee, S.; Lee, C.-G.; Lai, J. A novel framework for Bayesian calibration of building energy models with sub-hourly building operational data. In Proceedings of the 2023 Building Simulation Conference, Shanghai, China, 4–6 September 2023; pp. 1039–1046.
68. Xia, D.; Xie, W.; Guo, J.; Zou, Y.; Wu, Z.; Fan, Y. Building thermal and energy performance of subtropical terraced houses under future climate uncertainty. *Sustainability* **2023**, *15*, 12464. [[CrossRef](#)]
69. SNI 6389:2020; Energy Conservation in Building Envelopes. Badan Standardisasi Nasional (BSN): Jakarta, Indonesia, 2020.
70. IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013; p. 1535.
71. Taylor, K.E.; Stouffer, R.J.; Meehl, G.A. An Overview of CMIP5 and the Experiment Design. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 485–498. [[CrossRef](#)]
72. Remund, J.; Müller, S.; Schmutz, M.; Barsotti, D.; Graf, P.; Cattin, R. Meteonorm 8: Handbook Part I: Software, Version 8.2. 2023. Available online: https://mn8.meteonorm.com/assets/downloads/mn82_software.pdf (accessed on 28 December 2025).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.